Linear Algebra and its Applications 436 (2012) 2531–2535

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Remarks on the Böttcher–Wenzel inequality

Zhiqin Lu [1]

Department of Mathematics, University of California, Irvine, CA 92697, United States

ARTICLE INFO

Article history:
Received 5 June 2011
Accepted 12 September 2011
Available online 6 October 2011

Submitted by R.A. Brualdi

AMS classification:
Primary: 15A52
15A45
60H25
65F35

Keywords:
Commutator
Matrix inequalities
Best constants

ABSTRACT

In 2005, Böttcher and Wenzel raised the conjecture that if X, Y are real square matrices, then $||XY - YX||^2 \leqslant 2||X||^2||Y||^2$, where $||\cdot||$ is the Frobenius norm. Various proofs of this conjecture were found in the last few years by several authors. We here give another proof. This proof is highly conceptual and requires minimal computation. We also briefly discuss related inequalities, in particular, the classical Chern-do Camo–Kobayashi inequality.

1. The proof

The Böttcher–Wenzel inequality was proved in [2] for 2×2 matrices and by László [8] for 3×3 matrices. Different proofs of the full version of the conjecture were obtained by Lu [9], Vong and Jin [10], Böttcher–Wenzel [3], and Audenaert [1]. The complex matrix case was treated in Böttcher–Wenzel [3] and Wenzel [11]. A convenient observation that links the complex case to the real one can be found in Cheng–Vong–Wenzel [4, p. 296]. A useful observation was obtained in Audenaert [12] for further generalizations.

In this section, we give a new proof. Let $[X, Y]$ denote the commutator of X and Y and let $\|\cdot\|$ stand for the Frobenius norm (= Hilbert–Schmidt norm).

Theorem 1. *Let X, Y be real $n \times n$ matrices. Then*

$$||[X, Y]||^2 \leqslant 2||X||^2 \cdot ||Y||^2.$$

E-mail address: zlu@uci.edu

[1] The author is partially supported by NSF award DMS-0904653.

0024-3795/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2011.09.016

In accordance with [9, Lemma 3], we made the following definition. Let $V = \mathfrak{gl}(n, \mathbb{R})$ and define a linear map $T = T_X$ by

$$T : V \to V, \quad Y \mapsto [X^T, [X, Y]],$$

where X^T is the transpose of X.

Let

$$\Lambda = \begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}$$

and assume that

$$s_1^2 + \cdots + s_n^2 = 1. \tag{1}$$

Put

$$\Lambda_1 = \begin{pmatrix} 0 & \Lambda \\ 0 & 0 \end{pmatrix}.$$

Let $\tilde{T}_{\Lambda_1}$ be the restriction of T_{Λ_1} on $V = \mathfrak{gl}(n, \mathbb{R}) \oplus \mathfrak{gl}(n, \mathbb{R}) \subset \mathfrak{gl}(2n, \mathbb{R})$. Let finally

$$A = \begin{pmatrix} C & \\ & B \end{pmatrix}. \tag{2}$$

Then we have

$$[\Lambda_1, A] = \begin{pmatrix} 0 & \Lambda B - C\Lambda \\ 0 & 0 \end{pmatrix}, \tag{3}$$

and

$$\tilde{T}_{\Lambda_1}(A) = \begin{pmatrix} -(\Lambda B - C\Lambda)\Lambda & \\ & \Lambda(\Lambda B - C\Lambda) \end{pmatrix}.$$

For the rest of the paper, we make the following generic condition: all s_i are distinct and nonzero, $s_1^2 > s_2^2 > \cdots > s_n^2$, and all $s_i^2 + s_j^2$ are distinct.

Lemma 1. *The eigenvalues of $\tilde{T}_{\Lambda_1}$ must either be 0 or of the form $s_i^2 + s_j^2$. Let E_{ij} be the matrices whose only nonzero entry 1 is the (i, j)th entry. Then*

(1) The eigenspace of the eigenvalue $2s_i^2$ is spanned by $(B, C) = (E_{ii}, -E_{ii})$ for $1 \leqslant i \leqslant n$;
(2) The eigenspace of the eigenvalue $s_i^2 + s_j^2$ for $i \neq j$ is spanned by $(B, C) = (E_{ij}, -\frac{s_j}{s_i}E_{ij})$ and $(B, C) = (E_{ji}, -\frac{s_j}{s_i}E_{ji})$ for $1 \leqslant i \neq j \leqslant n$;
(3) The eigenspace of the eigenvalue 0 is spanned by $(B, C) = (E_{ij}, \frac{s_i}{s_j}E_{ij})$ for $1 \leqslant i, j \leqslant n$.

In particular, the maximum eigenvalue of $\tilde{T}_{\Lambda_1}$ is $2s_1^2$, of multiplicity 1, and the second largest eigenvalue of $\tilde{T}_{\Lambda_1}$ is $s_1^2 + s_2^2$, of multiplicity 2.

Proof. Let A in (2) be an eigenvector of the eigenvalue λ of $\tilde{T}_{\Lambda_1}$. Then we have

$$-(\Lambda B - C\Lambda)\Lambda = \lambda C, \quad \Lambda(\Lambda B - C\Lambda) = \lambda B.$$

Assuming that b_{ij}, c_{ij} are the entries of B, C, respectively, we have

$$-s_is_jb_{ij} = (\lambda - s_j^2)c_{ij}, \quad -s_is_jc_{ij} = (\lambda - s_i^2)b_{ij} \tag{4}$$

for $1 \leqslant i, j \leqslant n$. From the above equations, we conclude that the eigenvalues of $\tilde{T}_{\Lambda_1}$ must be the solutions of the equations

$$s_i^2s_j^2b_{ij}c_{ij} = (\lambda - s_i^2)(\lambda - s_j^2)b_{ij}c_{ij}$$

and hence be either 0 or $s_i^2 + s_j^2$. Moreover, for fixed (i, j) and the fixed eigenvalue $s_i^2 + s_j^2$, we have $b_{rs}c_{rs} = 0$ except $(r, s) = (i, j)$ or (j, i). Using this observation, we find all the eigenvectors of the operator $\tilde{T}_{\Lambda_1}$. $\square$

Proof of the Theroem 1. Following [2], we work with the singular value decomposition. Let $||X|| = 1$ and let

$$X = Q_1\Lambda Q_2 \tag{5}$$

be the singular decomposition of X, where Q_1, Q_2 are orthogonal matrices and Λ is a diagonal matrix. Let

$$B = Q_2YQ_2^{-1}, \quad C = Q_1^{-1}YQ_1.$$

Then we have

$$||[X, Y]||^2 = ||\Lambda B - C\Lambda||^2. \tag{6}$$

For fixed X, let Y be a matrix with unit norm such that $||[X, Y]||$ is maximized. Then we have

$$T_X(Y) = \lambda' Y$$

by the method of Lagrange multipliers. By [9, Proposition 5] (see also [4, Proposition 2.4]), $[X^T, Y^T]$ is also an eigenvector of λ' and it is linearly independent to Y. Let

$$Z = \alpha Y + \beta[X^T, Y^T]$$

be a linear combination of Y, $[X^T, Y^T]$ such that

$$A = \begin{pmatrix} Q_1^{-1}ZQ_1 & \\ & Q_2ZQ_2^{-1} \end{pmatrix}$$

is orthogonal to the first eigenspace of $\tilde{T}_{\Lambda_1}$. Since the space of all such A is 2-dimensional, the linear combination always exists. By (3), (6), we have

$$||[X, Y]||^2 = ||[\Lambda_1, A]||^2 = \langle A, \tilde{T}_{\Lambda_1}(A)\rangle.$$

Using Lemma 1, we have

$$||[X, Y]||^2 \leqslant (s_1^2 + s_2^2)||A||^2 \leqslant ||B||^2 + ||C||^2 = 2||Y||^2,$$

and the theorem is proved. $\square$

2. Additional remarks

Here are some remarks on further generalizations of the Böttcher–Wenzel inequality. We first prove the following result.

Theorem 2. *Let X, Y be $n \times n$ matrices. Let X be a diagonal matrix and let $||Y||_\infty = \max_{i \neq j}(|y_{ij}|)$, where (y_{ij}) are the entries of Y. Then we have*

$$||[X, Y]||^2 \leqslant ||X||^2 \cdot (||Y||^2 + 2||Y||^2_\infty). \tag{7}$$

Proof. In [9, pp. 1293, Remark 1], the theorem was proved assuming that Y is symmetric. That is, for any real numbers $\lambda_1, \ldots, \lambda_n$, we have

$$2\sum_{i<j}(\lambda_i - \lambda_j)^2 y_{ij}^2 \leqslant \left(\sum_j \lambda_j^2\right) \cdot \left(2\sum_{i>j} y_{ij}^2 + 2\max_{i>j}(y_{ij})^2\right).$$

This implies that, for strictly upper triangular matrix Y_1,

$$||[X, Y_1]||^2 \leqslant ||X||^2 \cdot (||Y_1||^2 + ||Y_1||^2_\infty).$$

By using the same argument, the above inequality is also true for strictly lower triangular matrices. Let $Y = Y_0 + Y_1 + Y_2$, where Y_0, Y_1, Y_2 are the diagonal part, the strictly upper triangular part, and the strictly lower triangular parts of Y, respectively. Then we have

$$\begin{aligned} ||[X, Y]||^2 &= ||[X, Y_1]||^2 + ||[X, Y_2]||^2 \\ &\leqslant ||X||^2 \cdot (||Y_1||^2 + ||Y_1||^2_\infty + ||Y_2||^2 + ||Y_2||^2_\infty) \\ &\leqslant ||X||^2 \cdot (||Y||^2 + 2||Y||^2_\infty), \end{aligned}$$

and the theorem is proved. □

Chern-do Carmo–Kobayashi [5] already had the Böttcher–Wenzel inequality when one of the matrices is symmetric. Their proof actually yields the following result.

Theorem 3. *Let X be a symmetric $n \times n$ matrix with λ_1 being the largest eigenvalue and λ_n being the smallest eigenvalue. Let Y be an $n \times n$ matrix. Then*

$$||[X, Y]||^2 \leqslant (\lambda_1 - \lambda_n)^2 ||Y||^2.$$

Proof. Without loss of generality, we may assume that X is a diagonal matrix. Thus we have

$$||[X, Y]||^2 = \sum_{i,j}(\lambda_i - \lambda_j)^2 y_{ij}^2 \leqslant (\lambda_1 - \lambda_n)^2 ||Y||^2,$$

where (y_{ij}) are the entries of Y. □

Let X be a real $n \times n$ matrix and let $s_1, \ldots, s_n$ be the singular values of X. The (2, (2))-Ky Fan norm of X is defined as

$$||X||_{2,(2)} = \sqrt{s_1^2 + s_2^2}.$$

Obviously, we have $||X||_{2,(2)} \leqslant ||X||$. From the proof of Theorem 1, we actually have

$$||[X, Y]||^2 \leqslant 2||X||^2_{2,(2)}||Y||^2, \tag{8}$$

which is a generalization of the Böttcher–Wenzel inequality. This strengthened inequality was first proved by Wenzel [11].

Evidently, we have

$$(\lambda_1 - \lambda_n)^2 \leqslant 2 \max_{i \neq j} (\lambda_i^2 + \lambda_j^2).$$

Therefore, if X is a symmetric matrix, then the Chern-do Carmo–Kobayashi inequality is sharper than the Wenzel inequality (8). On the other hand, in a lot of cases inequality (7) is sharper than both the Chern-do Carmo–Kobayashi and the Wenzel's inequalities because the ∞ norm is usually much smaller. We wish to obtain a common generalization of the above three inequalities. Such a result would provide a common generalization of the Böttcher–Wenzel inequality and the Normal Scalar Curvature inequality proved in [9,7].

Finally, a generalization of Theorem 2 may exist in connection with the p Schatten norms considered in [4,11,12].

Acknowledgement

The author deeply thanks Professors A. Böttcher and D. Wenzel for their many useful comments without which the paper would not be in its current form.

References

[1] K.M.R. Audenaert, Variance bounds, with an application to norm bounds for commutators, Linear Algebra Appl. 432 (5) (2010) 1126–1143, doi:10.1016/j.laa.2009.10.022, MR 2577614 (2011b:15047).
[2] A. Böttcher, D. Wenzel, How big can the commutator of two matrices be and how big is it typically?, Linear Algebra Appl. 403 (2005) 216–228, doi:10.1016/j.laa.2005.02.012, MR 2140283 (2006f:15016).
[3] A. Böttcher, D. Wenzel, The Frobenius norm and the commutator, Linear Algebra Appl. 429 (8–9) (2008) 1864–1885, doi:10.1016/j.laa.2008.05.020, MR 2446625 (2009j:15089).
[4] C-M. Cheng, S-W. Vong, D. Wenzel, Commutators with maximal Frobenius norm, Linear Algebra Appl. 432 (1) (2010) 292–306, doi:10.1016/j.laa.2009.08.008, MR 2566477 (2010m:15031).
[5] S.S. Chern, M. do Carmo, S. Kobayashi, Minimal submanifolds of a sphere with second fundamental form of constant length, in: Functional Analysis and Related Fields, Proc. Conf. for M. Stone, Univ. Chicago, Chicago, Ill., 1968, Springer, New York (1970) 59–75, MR0273546 (42 #8424).
[6] K-S. Fong, C-M. Cheng, I-K. Lok, Another unitarily invariant norm attaining the minimum norm bound for commutators, Linear Algebra Appl. 433 (11–12) (2010) 1793–1797, doi:10.1016/j.laa.2010.06.037, MR 2736098.
[7] J. Ge, Z. Tang, A proof of the DDVV conjecture and its equality case, Pacific J. Math. 237 (1) (2008) 87–95, doi:10.2140/pjm.2008.237.87, MR2415209 (2009d:53080).
[8] L. László, Proof of Böttcher and Wenzel's conjecture on commutator norms for 3-by-3 matrices, Linear Algebra Appl. 422 (2–3) (2007) 659–663, doi:10.1016/j.laa.2006.11.021, MR 2305148 (2008a:15053).
[9] Z. Lu, Normal scalar curvature conjecture and its applications, J. Funct. Anal. 261 (2011) 1284–1308, doi:10.1016/j.jfa.2011.05.002.
[10] S-W. Vong, X-Q. Jin, Proof of Böttcher and Wenzel's conjecture, Oper. Matrices 2 (3) (2008) 435–442, MR 2440678 (2009f:15033).
[11] D. Wenzel, Dominating the commutator, Oper. Theory Adv. Appl. vol. 202, Birkhäuser Verlag, Basel (2010) 579–600, MP 2723302 (2011j:15034).
[12] D. Wenzel, K.M.R. Audenaert, Impressions of convexity: an illustration for commutator bounds, Linear Algebra Appl. 433 (11–12) (2010) 1726–1759, doi:10.1016/j.laa.2010.06.039, MR 2736095.